R E P O R T R E S U M E S

ED 002 974 24

PROBLEM-SOLVING PROFICIENCY AMONG ELEMENTARY SCHOOL TEACHERS, IV. FURTHER INVESTIGATIONS OF TEACHERS OF ARITHMETIC, GRADES 3-6.

BY- TURNER, RICHARD L.
INDIANA UNIV., BLOOMINGTON, INST. OF EDUC. RES.

REPORT NUMBER CRP-419-4 PUB DATE FEB 61
CONTRACT OEC-7790
EDRS PRICE MF-$0.09 HC-$2.08 52P.

DESCRIPTORS- *ARITHMETIC, *PROBLEM SOLVING, *ELEMENTARY SCHOOL TEACHERS, TEACHER CHARACTERISTICS, EFFECTIVE TEACHING, *TEACHING TECHNIQUES, INTELLIGENCE, *TEACHER EVALUATION, BLOOMINGTON, INDIANA, ALLPORT VERNON LINDZEY STUDY OF VALUES, MINNESOTA TEACHER ATTITUDE INVENTORY (MTAI)

THE QUESTION OF THIS STUDY WAS WHETHER PREVIOUSLY FOUND RELATIONSHIPS BETWEEN INTELLIGENCE AND PROBLEM SOLVING WOULD HOLD UP EVEN THOUGH SEVERAL MODIFICATIONS HAD BEEN INTRODUCED INTO THE CRITERION INSTRUMENT. SAMPLES WERE DRAWN FROM BOTH PROFESSIONAL UNDERGRADUATE AND GRADUATE COURSES CONCERNED WITH THE TEACHING OF ARITHMETIC AT INDIANA UNIVERSITY AND FROM SEVERAL INDIANA SCHOOL SYSTEMS. BOTH WERE GIVEN THE CRITERION INSTRUMENT, THE MINNESOTA TEACHER ATTITUDE INVENTORY (MTAI), AND THE "STUDY OF VALUES." FROM THE RESULTS, IT WAS APPARENT THAT THOSE HYPOTHESES STEMMING FROM THE MTAI AND "STUDY OF VALUES" WERE NOT CONFIRMED. THE BELIEF THAT PERFORMANCE VARIES WITH OPPORTUNITY TO LEARN NEEDS FURTHER CAREFUL EXAMINATION. IT IS POSSIBLE THAT A COMPLEX SET OF RELATIONSHIPS EXISTS BETWEEN OPPORTUNITIES TO LEARN, ABILITY TO LEARN, AND TEACHER SELECTION PRACTICES IN SCHOOL SYSTEMS. WHILE SUCH RELATIONSHIPS COULD NOT BE EXTRACTED FROM THE PRESENT DATA, THE INVESTIGATOR FELT THAT THE VERY PROBABILITY THAT THEY EXIST SHOULD NOT BE OVERLOOKED IN SUBSEQUENT STUDIES. (GD)

ED002974

REFERENCE
COPY

COOPERATIVE
RESEARCH
PROJECT

No. 419-IV

419

MONOGRAPH
of the
INSTITUTE OF EDUCATIONAL RESEARCH
at
INDIANA UNIVERSITY

Problem Solving Proficiency Among Elementary School Teachers

IV. Further Investigations of Teachers of Arithmetic Grades 3-6

by
Richard L. Turner

FEBRUARY 1961

PROBLEM SOLVING PROFICIENCY AMONG ELEMENTARY SCHOOL TEACHERS

IV. FURTHER INVESTIGATIONS OF TEACHERS OF ARITHMETIC GRADES 3-6

by

Richard L. Turner

The research on which this study was based was supported in part by the Cooperative Research Program of the U. S. Office of Education under Project 419 (7790).

Published by the
INSTITUTE OF EDUCATIONAL RESEARCH
School of Education
Indiana University

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

Price, $1.00, postpaid
For sale by the Institute of Educational Research
School of Education
Bloomington, Indiana

TABLE OF CONTENTS

	Page
INTRODUCTION	1
PROCEDURES	9
Problem Construction	9
Sampling	17
RESULTS	21
Variables Bearing on Undergraduate Performance	21
Variables Bearing on Teacher Performance	26
Interproblem Relationships	37
DISCUSSION AND CONCLUSIONS	40

LIST OF TABLES

Table		Page
1.	Means and Variances of Pre-test and Post-test Scores . . .	22
2.	Analysis of Variance of Pre-test Scores among Undergraduates	23
3.	Correlations Between ACE, Coop C_2 and Step Scores and the Pre-test, Post-test, and Gains of Student Teachers .	25
4.	Correlations Between MTAI and Study of Values Scores, and the Post-test Scores on Student Teachers	26
5.	Performance of Teachers of Grade 3 Versus Performance of Teachers of Grades 4-6	27
6.	Pre-test and Post-test Means of Student Teachers of Grade 3 and Student Teachers of Grades 4-6	28
7.	Means and Variances of Teachers Grades 4-6 Public Versus Private Colleges, Small Versus Large Cities	29
8.	Analysis of Variance City Size X Type of Teacher Preparatory Institution	30
9.	Means and Variances of Teachers with 0, 1-2, 3-5, and 6-10 Years of Teaching Experience Prepared at Public Institutions	31
10.	Means and Variances of Teachers with 0 Years of Experience Prepared at a Public Institution and 1-2, 3-5, and 6-10 Years Experience Prepared at Private Institutions . . .	32
11.	Means and Variances of Teachers with 0, 1-2, 3-5, and 6-10 Years of Teaching Experience Teaching in Cities 25,000 and Above	33
12.	Means and Variances of Teachers with 0 Years of Experimental Teaching in Cities 25,000 and Above, and Teachers With 1-2, 3-5, and 6-10 Years of Experience Teaching in Cities 24,999 and Below	34

Table | Page

13. Teachers of Grades 4-6 Versus Teachers of Grades K-3 at Four Levels of Teaching Experience 35

14. Means and Variances of Teachers 20-29, 30-39, and 40+ Years of Age 36

15. Correlations Between MTAI Scores, Study of Values Scales, and the Problem Solving Performance of Teachers Grades 4-6 37

16. Interproblem relationships for 69 Teachers Grades 4-6 . . . 38

17. Interproblem relationships for 47 Student Teachers Grades 3-6. 38

INTRODUCTION

In an earlier statement of research strategy with respect to problem solving proficiency among elementary school teachers, Turner and Fattu (7) asserted a number of hypotheses about the independent variables that would be expected to be related to individual differences in ability to solve standard problems in teaching elementary school. The purpose of the present study is to examine these hypotheses, with particular attention to the specific conditions under which they may be verified for teachers of arithmetic in grades 3-6.

The specific hypotheses of interest in the present study are those earlier generated by use of the assumption that problem solving proficiency among elementary teachers is contingent upon the array of instrumental responses these teachers have acquired. One of these hypotheses (7, p. 16) is that measured intelligence is positively related to the acquisition of instrumental responses relevant to teaching and should thus be related to performance on problems in which these responses are utilized. That there is a relationship between problem performance and intelligence has previously been verified with respect to problem solving proficiency in teaching arithmetic (9) and problem solving proficiency in teaching reading (10). The question in the present study was whether the previously found relationships between intelligence and problem solving proficiency in arithmetic would recur even though several modifications had been introduced into the criterion instrument.

Several hypothesis generated on the assumption that proficiency is contingent upon the array of instrumental responses available to a teacher

bear on differences in the opportunity to acquire such responses. One of these hypotheses is that there is differential acquisition of responses according to the institution at which a teacher is prepared (7, p. 17). This hypothesis may be tested by examining teacher performance according to preparatory source. Turner (9) has previously verified that teachers do in fact differ according to source of preparation. There is however, a further implication in this hypothesis. If teachers vary according to the institution at which they were prepared, and this variation is truly attributable to differences in learning, then it should be possible to show that there are some learning treatments which modify performance, within an institution in that group of institutions with which higher performance is associated. While there might be any number of treatments during teacher preparation that are relevant to problem solving performance in teaching arithmetic, some of these, namely courses in methods of teaching arithmetic, and student teaching, are conceptually more relevant than others.

In his previous investigation of the influence of a course in methods of teaching arithmetic and of student teaching on proficiency, Turner (9) failed to find significant results. Aside from the possibility that these treatments may actually not be associated with changes in proficiency, two additional factors in the failure to get results should be considered. One is, that the instrument used to assess proficiency was comparatively unrefined, never having been item analyzed. A second factor is that matched pairs were used to determine whether the arithmetic methods course contributed to differences in performance. Students just entering the course were matched on age, sex, and ACE score to those just completing the course. Since neither age nor sex

were shown to be significant variables in performance and intelligence accounted for not more than 20% of the variance, whether matching students on these variables represented adequate control is doubtful.

In the present study the hypothesis that teachers vary in performance according to preparatory institution and the hypothesis that an arithmetic methods course, within an institution within that group of institutions with which higher performance is associated, and student teaching are significantly related to performance were both tested again, with somewhat better controls exerted in testing the latter hypothesis than was previously the case.

In addition to the two foregoing hypotheses, there are two others which stem from the notion that differences in the opportunity to acquire instrumental responses affect performance. The first of these is that the type of situation in which the teacher is employed is a variable in problem solving proficiency among teachers (7, p. 17). This variable is hypothesized to be significant because the opportunity to acquire instrumental responses in a particular teaching situation is thought to vary with the amount of autonomy a teacher has, such that the wider the scope of responses permitted the teacher in the classroom, the greater the opportunity to acquire a full range of responses relevant to solving problems in teaching. However, it is apparent that whether a teacher can take advantage of broad autonomy depends partly on the extent to which he is prepared to do so by previous education, particularly whether he has been prepared to try to bring about a full range of objectives and intelligently experiment with a wide range of techniques in bringing about these objectives. The latter notion suggests, then, that is is not autonomy alone which would

affect performance, but an interaction between autonomy and preparation.

The direct testing of the effects of autonomy and the effects of preparation and the interaction between would require independent assessments of both autonomy and of preparatory programs. To assess these variables by independent measures was considered an unfeasible undertaking since the assessing devices would first have to be constructed and validated. It was thus necessary to begin, in the present study, by testing the hypotheses that both the preparatory source of teachers and the type of situation in which they teach, as indexed by the size of the school system, and the interaction between would be significant sources of variation in performance.

It should be made clear that the hypotheses tested by no means necessarily demonstrate differences in autonomy between situations or genuine differences in preparation; these hypotheses bear only on whether it is reasonable to continue to suppose that autonomy and preparation are significant variables in performance. Demonstration that these variables are definitely sources of variations presupposes independent assessment of them.

In a previous study, Turner tested for differences between county, town and city teachers drawn from two consolidated systems, and for differences between the two consolidated systems as such, and found no significant results. The failure to get results in the latter case seems largely attributable to having sampled systems reasonably homogeneous with respect to administrative practice and degree of teacher control. However, because the central purpose of the previous study was to demonstrate differences between a representative cross-section of teachers and a group of non-teachers, as a

point of problem validation, to draw samples from extremely divergent systems to demonstrate differences between them would have been inappropriate.

A final hypothesis stemming from the notion that differences in opportunities to learn are reflected in performance is that problem solving proficiency is a function of teaching experience. This hypothesis, however, needs very careful qualifications. Both type of preparation and type of teaching situation and the interaction between have been hypothesized to bear on problem solving performance. Since teaching experience must occur after some type of preparation and in particular situations, the specific relationships between experience and performance would be expected to be contingent upon these variables. In other words, an accurate statement about the effects of teaching experience on problem solving performance must take into account both the preparatory antecedents to this experience and the conditions under which the experience has occured.

In a previous study, Turner found that whether cross-sectional groups of teachers showed increasingly better performance as a function of years of experience appeared to be contingent upon the type of institution at which the teacher was prepared. However, because of an insufficient number of subjects in some of the sub-groups, this finding may be considered only suggestive. In the present study changes in performance with experience both according to preparatory antecedents and according to type of teaching situation were investigated, in part to clarify these previous results.

In addition to the above hypotheses which stem from the assumption that the performance varies according to opportunity to acquire instrumental

responses, other hypotheses, stemming from the assumption that response acquisition is a function of the motivation of the teacher, may also be made. With respect to teacher motivation and problem solving performance, there are essentially two different sets of questions. One set deals with the type of events that can be shown to increase the probability of a given teacher response, i.e. reinforcing events. The answers to this set of questions must presumably be found through experimentation. The second set of questions deals with whether certain types of measured values, motives, attitudes or interests are associated with proficiency in problem solving. The concern in the present report was with the latter kinds of questions.

In a previous paper, Turner found that among undergraduates, teacher attitudes as measured by the MTAI were positively related to problem solving performance in arithmetic. This result, it should be noted, held up under cross-validation in the same paper. The questions investigated in the present study in the latter respect were: (i) whether the same results would hold for a third undergraduate sample; and (ii) whether MTAI scores are positively associated with problem performance among experienced teachers. Turner also previously investigated relationships between dominant interests as measured by the Study of Values and problem solving performance. The results were inconsistent among undergraduates; the initial group showed positive relationships between the aesthetic and the social scales and performance, while the cross-validating group failed to show positive relationships on these scales, but showed a positive relationship between the religious scale and performance. These results indicated that the relationships between the

separate scales on the Study of Values and problem solving performance is tenuous, at least among undergraduates. The questions investigated in the present study with respect to the Study of Values were: (i) whether the relationships between the various scales and performance among undergraduates would show any greater consistency than previously, and (ii) whether there are consistent relationships between the problem solving performance and the individual scales among experienced teachers.

The hypotheses and questions stated above were those asserted or implied by Turner and Fattu in earlier statements of research strategy. However, in addition to these hypotheses and questions, there are a number of questions that stem primarily from the nature of the instrument used to assess problem solving proficiency in arithmetic. One of these questions is whether the problems used in the instrument are valid for the full grade range 3-6. While Turner previously found no differences in performance by grade level, the instrument used by him was modified for the present study, and one of the problems previously directed to grade 3 completely deleted. Whether the modifications made, or possibly other variables, affect the relevance of the problems for teachers at particular grade levels was clearly a focal question.

A second question, relevant primarily to the arithmetic context of the problems, is the extent to which performance varies with arithmetic achievement. In the preceeding study, Turner found no relationship between either number of courses in mathematics (up to eight such courses, including high school courses) or recency of exposure to arithmetic methods and problem performance. However, these variables are quite rough indices of arithmetic

achievement. In the present study, arithmetic achievement as measured by a standardized test replaces them.

A third question, relevant primarily because the problems describing the criterion measure must be read, is whether reading comprehension is related to problem performance. Turner has previously shown that reading comprehension correlated from .29 to .50 with performance among undergraduates. However, since intelligence as measured by the ACE correlated between .74 and .84 with reading comprehension, it seems probable that reading comprehension alone accounts for very little of the variance of performance. The question in the present study was simply whether the positive relationship between reading comprehension and performance would continue to appear.

A final variable, age, is not of particular interest in and of itself, in relation to performance, but of interest because it is confounded with the length of teaching experience. In the previous study age was shown to be of no significance up to age 50, with experience held constant from 1-7 years. The question in the present study was whether this same general result would obtain, thus eliminating concern over whether changes in performance are a function of age as well as experience.

PROCEDURES

Problem Construction

To obtain the criterion measures of problem solving proficiency in teaching arithmetic, the seven problems (Form D) previously described and used by Turner (9) were modified into a similar but new instrument, and then a conceptually parallel form to this new instrument constructed, producing Forms E and F respectively.

To construct Form E, Chi Square values for each of the sixty-nine scored response alternatives in the seven problems in Form D were computed. Two groups of teachers, one from a large consolidated school system and one from a "small" consolidated system, student teachers, and non-teachers were used as criterion groups. In addition, the number of teachers correctly responding to each item on each problem was observed as an index of item difficulty. Use of these procedures indicated that one of the problems was vegetative, i.e. contained no items that discriminated among the criterion groups. The problem was thus deleted. A second problem was found to be an excellant discriminator but too difficult for all but a few teachers. It was deleted from Form E, but subsequently revised and placed in Form F.

The remaining five of the original seven problems were retained, with various modifications, in Form E. Of these five, three required changes in instructions to produce greater clarity; in addition the method of responding to two of these three was changed: With respect to the remaining three problems, one was left unchanged and in the other only those response alternatives

which failed to discriminate among criterion groups were deleted.

The five problems in Form F were constructed after the above revisions were made in the five problems for Form E. The construction of Form F was undertaken, of course, to establish an alternate form to E. The extent to which each problem in Form F is parallel to its mate in Form E may be observed in the problem descriptions and observations of equivalence given below.

Problem 1, Form E. This problem is based on problem H of Form D (9, p. 12) in which the subject was asked to identify the degree of relevance of 10 stated objectives of arithmetic instruction to six problems composing an arithmetic exercise. The central differences between problem H, Form D and Problem 1, Form E lie in a change in the method of responding to the Problem and change of instructions. In problem H the subject was required to circle for each of ten alternative objectives whether the objective was: A. relevant to the exercise (the exercise could definitely be used to help attain that objective); B. moderately relevant (the exercise might or could possibly be used to help attain that objective); C. not relevant (the exercise could not be used to attain that objective). In Problem 1, Form E (Problem 1-E) the subject is required to place the number of each alternative objective in one of four concentric circles, depending on how closely related he believes the objective is to the six problems given in the exercise, such that the most closely related objectives are placed in the center circle or "bull's eye" and the least closely related in the outermost circle, with the remaining objectives ordered between according to the closeness of the relationship. This problem has a maximum total score of 10 points, one for each alternative correctly ordered.

Problem 1, Form F. This problem utilizes identical objectives to those stated in Problem 1-E and is responded to by the same method, but 7 problems differing in content and structure from those in 1-E are presented in the exercise. Consequently, the relationships between the objectives and the exercise change in comparison to Problem 1-E, although the problem for the subject remains, conceptually speaking, the same.

Equivalence of Problems 1-E and 1-F. To determine the equivalence of problems 1-E and 1-F, the two problems were administered to a heterogeneous group of 22 graduate students in education such that half the group did 1-E first and 1-F second and the other half 1-F first and 1-E second. The r between the two problems was .414. The mean score on problem 1-F was 3.40; on 1-F, 2.77; the variance for 1-E, 2.63; for 1-F, 1.72; neither the means nor the variances differed significantly. The failure of these two problems to correlate highly is probably attributable to too great a shift in the arithmetic problems constituting the exercise. Problem 1-E and 1-F were used as described in the present study, since some data had been collected when the non-equivalence of these two problems was discovered. However, Problem 1-F has subsequently been revised.

Problem 2, Form E. This problem is based on problem A, Form D, (9, p. 11) in which the subject was asked to examine 10 examples in long division worked by a "pupil" in grade five. In these examples, the "pupil" systematically errors in solving one class of division problems and errors randomly in two other classes. The problem for the teacher was to identify which combination of 14 alternative teacher actions would be necessary and sufficient to

correct the systematic errors. In Problem 2-E, the examples and the errors in working the examples made by the pupil remain identical to those in Problem A, Form D; but the problem for the subject was changed from identifying which combination of 14 teacher actions are necessary and sufficient, to identifying which of ten statements of error are: (i) most closely identified with the systematic errors of the pupil and should thus be made the focus of an interview with the pupil; (ii) which statements or error are peripheral, but could be included in the interview; and (iii) which statements of error have no bearing on the errors made by the pupil.

Problem 2, Form F. This problem is structurally identical to Problem 2-E and utilizes the same response alternatives; however, the long division problems worked by the pupil involve systematic errors in a slightly different class of examples than in Problem 2-E.

Equivalence of Problems 2-E and 2-F. To determine the equivalence of Problems 2-E and 2-F, the two problems were administered to a heterogeneous group of 21 graduate students in education such that approximately half the group did Problem 2-E first and 2-F second and the remainder 2-F first and 2-E second. The r between the two problems was .371; the mean of 2-E was 5.57; of 2-F, 4.81; the variance of 2-E, 1.63; of 2-F, .45. There was no significant difference between the means, but the variances differ at p = .02. Further examination of the differences indicated that the group which took 2-F first had a significantly lower mean (p = .05) than the group that took 2-E first, whereas mean performance on the problem taken second did not differ significantly. This finding was interpreted to mean that 2-F is a significantly more

difficult problem when encountered without practice than is 2-E, a startling finding when the slight differences between the two problems are considered.* The problems were used as described in the present study, since some data had already been collected. However, Problem 2-F has subsequently been changed.

Problem 3, Form E. This problem is based on Problem E, Form D, (9, p. 12) in which the subject was asked to examine 12 mixed examples and six mixed problems worked by a pupil in grade 4. As in Problem 2-F, the errors made by the pupil were systematic in one class of problems and random in two others, and the subject was asked to determine which of 15 statements of error were most closely identified with the systematic errors of the pupil and should thus be made the focus of the interview, and so on as described in Problem 2-E above. The only changes made between Problem E, Form D and Problem 3-E was the elimination of 5 of the distractor response alternatives which were indicated by item analysis to be gratuitous.

Problem 3, Form F. The structure and response alternatives in this problem are identical to Problem 3-E. The examples and problems to be examined by the teachers as the basis for responding are members of the same classes as those in Problem 3-E, with changes only in numbers and, in the case of story problems, in wording.

Equivalence of Problems 3-E and 3-F. To determine the equivalence of Problems 3-E and 3-F, the two problems were administered to a heterogeneous group of 24 graduate students in education such that half the group did

---

*The two problems involved may be inspected by the reader upon written request to the author.

Problem 3-E first and 3-F second and the remainder 3-F first and 3-E second. The r between the problems was .663; the mean for 3-E, 5.71; for 3-F, 5.83; the variance for 3-E, 3.52; for 3-F, 3.27. Neither the means nor the variances differed significantly.

Problem 4, Form E. This problem is identical to Problem F, Form D (9, p. 13), in which the subjects were asked to place seven examples in long division in rank order according to their difficulty level.

Problem 4, Form F. The structure of this problem is identical to that of Problem 4-E, only the numbers in the examples to be ranked are changed. It is crucial in this type of problem, of course, that the class to which an example belongs not be changed between forms of a problem, since such changes modify difficulty level.

Equivalence of Problems 4-E and 4-F. Problems 4-E and 4-F were administered to a heterogeneous group of 27 graduate students in education by the usual procedure. The r between problems was .791; the mean of 4-E, 6.00; of 4-F, 6.00; the variance of 4-E, 14.15; of 4-F, 17.54. Neither the means nor variances differed significantly.

Problem 5, Form E. This problem is based on Problem D, Form D (9, p. 13), in which the subject was asked to state the "meanings" of subtraction as generalizations by which pupils could determine whether or not subtraction would be useful in solving a problem. The major change between Problem D, Form D and Problem 5-E lies in asking the subject to state, in Problem 5-E, what he would consider good answers from a pupil answering the question "what are the questions that you could ask yourself to help you

decide whether or not subtraction is the right thing to do to solve a problem?"

Problem 5, Form F. This problem is based on Problem G, Form D (9, p.13) in which the subject was asked to state the "meanings" of division as generalizations by which pupils could determine whether or not division would be useful in solving a problem. The major change between Problem G, Form D and Problem 5-F lies in asking the subject to state, in Problem 5-F, what he would consider good answers from a pupil answering the question "what are the questions that you could ask yourself to help you decide whether or not division is the right thing to do to solve a problem?"

Equivalence of Problems 5-E and 5-F. The problems were administered to a heterogeneous group of 20 graduate students in education, according to the usual procedure. The r between problems was .593; the mean of 5-E, 2.4; of 5-F, 1.8; the variance of 5-E, 5.73; or 5-F, 3.96. Neither the means nor the variances differed significantly.

Estimates of the Reliability of Form 5-E and 5-F. The estimates of the reliability of Forms E and F were made by correcting the r between each pair of problems, as described above, for a test 5 times as long. This procedure yields a corrected coefficient of reliability of .78 when the estimate is based on Problems 1-E, 1-F; of .75 based on Problems 2-E, 2-F; of .90 based on Problems 3-E, 3-F; of .94 based on Problems 4-E, 4-F; and of .88 based on Probiems 5-E and 5-F. The lower bound of the reliability coefficient by these estimates would be .75, the upper bound .94, with a median coefficient of .88 and a mean coefficient of .85. Probably the latter coefficients are the best estimates of the reliability of the total score on each form. It

may be noted in this connection that the corrected split-half reliability for total score on Form D, from which the problems were derived, was .83 (9, p. 51).

Scoring changes between Form D and Forms E and F. In Form D, four score criteria were available: (i) the "PI score" or number of correctly identified goal achieving responses; (ii) the "VN score" or number of correctly eliminated distractors; (iii) the "CR score", i.e. PI + VN; and (iv) the "consistency criterion" or consistency with which the subject scored above random response level on both the PI and VN score on each problem. Of these score criteria, the PI score and the consistency criterion were shown to be discriminatory with respect to the criterion groups, while the VN score was non-discriminating, correspondingly reducing the power of the CR score.

Because it is statistically tedious to deal with four score criteria (dependent variables) in examining each independent variable, the problems in Forms E and F were constructed to yield a single, continuous score, highly sensitive to the consistency of performance and ability to eliminate distractors. In order to make a single score sensitive to the ability to eliminate distractors, while at the same time not destroying its validity, the least discriminating distractors were eliminated from each problem in which they had appeared in Form D, viz., Problems A (now 2-E), E (now 3-E), and H (now 1-E). With the elimination of the gratuitous distractors, the goal achieving responses and the remaining distractors were given equal score weight, namely one point each, yielding a total score of ten for each of Problems 1-E, 2-E, and 3-E. This move, however, reduced the sensitivity of the total continuous score with respect to ability to identify goal achieving responses, since two of the problems,

F (now 4-E), and D (now 5-E), yielded only responses counted as goal-achieving and were worth only seven and four points respectively. To correct this state of affairs, scores on the latter two problems were multiplied by 2, bringing their representation in the scores on Forms E and F into proportion with their representation on Form D.

The data with respect to the equivalence of each problem in Form E to its mate in Form F and the estimated reliability of these Forms are based on the single continuous score as detailed above.

Sampling

In addition to the samples reported in establishing the equivalence of the problems, and used only for that purpose, samples were drawn: (i) from professional undergraduate and graduate courses concerned with the teaching of arithmetic at Indiana University, and (ii) from several Indiana school systems.

Undergraduate Sample. All students enrolled in an undergraduate course in methods of teaching arithmetic and present in class for the second session of the third week of the first semester were assessed on either Form E or Form F by means of passing alternative Forms down alternate columns of chairs such that the first person in column 1 received Form E and the second person Form F, etc., while the first person in column 2 received Form F and the second person Form E, etc. One week later, the same students were assessed on the STEP, Form 3A Mathematics (4).

Approximately four months later, after completion of arithmetic methods, these students did student teaching in some grade between Kindergarten and grade 6 for a period of approximately 12 weeks, and then returned to campus for a workshop. All students present at the workshop on the fifth day before the end of the semester were again assessed, all receiving Form E of the arithmetic problems, which were completed during the session. In addition, the students were requested to complete at home the MTAI (2) and the Study of Values (1) to be returned the next day.

This procedure yielded 77 paired observations on the arithmetic problems, and 12 observations of students who had either been absent for the initial assessment or had completed the methods course elsewhere. Of the 77 paired observations, 16 students were pre-tested on Form F and post-tested on Form E and did student teaching in one of grades K-2. These students are designated group FEK-2. Nineteen of the 77 were pre-tested on Form F and post-tested on Form E, and did student teaching in one of grades 3-6. These students are designated FE3-6. Twenty-two students were both pre- and post-tested on Form E, and did student teaching in K-2. These students are designated EEK-2. Twenty students were both pre- and post-tested on Form E, and did student teaching in grades 3-6. These students are designated EE3-6. The unequal n's in these groups are attributable to differential losses from them in the interval between pre- and post-test. In addition to scores on the problems in teaching arithmetic, scores on the STEP, MTAI, Study of Values, and on the ACE Psychological Examination (6) and Coop C_2 were also obtained for undergraduates. The latter two scores were obtained from the University Test

Bureau, and represent performance on these instruments at the time of admission to the University. These scores were not available for all students.

Teacher Sample. The teacher sample was drawn from five different sources. The first source was a non-consolidated school system in a city of about 10,000 population, in which there were 12 teachers with 1-10 years experience teaching in grades 3-6. Nine of these teachers participated. The second source was a non-consolidated school system in a city of about 18,000 population in which there were 21 teachers with 1-10 years experience teaching in grades 3-6. All of these teachers participated. The third source was a consolidated schoolsystem centered in a city of about 40,000 population. All teachers grades 3-6 in this system had been assessed on Form D one year before, hence only those teachers with 1-10 years experience teaching in grades 3-6 who were new to the system were invited to participate. There were 22 such teachers, of whom 16 participated. The fourth source was a township school system in a city of about 500,000. Through a misunderstanding over whether beginning teachers in the system were to participate, only teachers with 2-10 years experience teaching in grades 3-6 were notified on the possibility of participation. Of the 44 such teachers in the system, 32 participated. The fifth source was a heterogeneous group of graduate students enrolled in a couse in improving arithmetic instruction. All students present for the third session of the second week of summer school were measured. Of this total group, only two subgroups were utilized in the present study: (i) 23 Indiana teachers with 1-10 years experience teaching in grades 3-6, and (ii) Indiana teachers with 1-10 years experience teaching in grades K-2. The latter group was used only in one

comparison in which the performance of non-intermediate grade teachers is focal.

The procedures used with participating teachers in school systems were as follows: All participating teachers met as a group, at which time they were administered Form E of the arithmetic problems. Upon completion of the problems, each participant received an envelope containing a questionnaire, the MTAI and the Study of Values, to be completed at home and returned by mail. In the case of the teachers enrolled in the graduate course in teaching arithmetic, the questionnaire, the MTAI, and Study of Values were returned to the instructor of the course rather than by mail.

RESULTS

Variables Bearing On Undergraduate Performance

Each of the variables investigated with respect to undergraduate performance stand in one or more of three relationships to this performance: 1. Exposure to arithmetic methods and subsequent student teaching in either grades K-2 or grades 3-6 are essentially treatment variables; 2. ACE Scores and Coop C_2 Scores stand in a predictor relationship to pre-test performance, gains and post-test performance, since these two tests were administered three years before the pre-test assessment of performance. The STEP stands as a predictor variable in relation to gains in performance and post-test performance, but as a concurrent variable with respect to pre-test performance, since it was administered at approximately the same time as the pre-test; 3. MTAI scores and Study of Values scores stand as concurrent variables with respect to post-test performance.

Arithmetic methods and student teaching. In examining these variables several matters need to be taken into account. First, while all students were post-tested on Form E, some were pre-tested on Form F. Thus the Form of the test taken is a variable to be considered. Second, while all students were, hypothetically speaking, exposed to the same treatment in the arithmetic methods course, student teaching treatments varied at least by grade level if not also in other ways. Conceptually, student teaching in any of grades K-2 is less relevant to performance on the problems in the criterion instrument

than is student teaching in any of grades 3-6.

The basic data with respect to each undergraduate group (EEK-2, EE3-6, FEK-2, and FE3-6) may be ovserved in Table 1. The initial question with respect to this data is whether the pre-test performance of these groups is homogeneous with respect to means and variances, thus indicating whether the form of the test taken in the pre-test is a significant source of variation. The answer to this question may be observed in Table 2.

TABLE 1. MEANS AND VARIANCES OF PRE-TEST AND POST-TEST SCORES.

Group				Post-test	
	n	$\bar{X}$	s^2	$\bar{X}$	s^2
EE K-2	22	23.18	47.24	24.55	49.29
FE K-2	16	17.75	52.60	24.43	25.60
EE 3-6	20	21.10	52.63	28.30	37.58
FE 3-6	19	19.21	46.01	27.21	21.28

On the basis of Tables 1 and 2, it is clear that there are no significant differences in pre-test performance according to test form, nor are there differences in means and variances between the K-2 and 3-6 groups. The latter observation may be confirmed by examining the pooled pre-test means for each of these groups with test form disregarded. The pre-test mean of all student teachers who taught K-2 is 20.92 and the variance 54.14, the pre-test mean of all student teachers who taught grades 3-6 is 20.18, and the variance 49.07.

Neither the means nor the variance of these pooled groups differs significantly.

TABLE 2. ANALYSIS OF VARIANCE OF PRE-TEST SCORES AMONG UNDERGRADUATES.

Sources of Variation	df	ss	Mean Square
Groups	3	318	106.00
Within	73	3609	49.44
Totals	76	3927	

Groups ÷ Within = 2.14, $p > .05$
F extreme variances = 1.14, $p > .05$

Examination of the differences between pre-test and post-test means in each group shown in Table 1 by means of the t test for correlated groups indicates that groups FEK-2, EE3-6 and FE3-6 all made significant gains (beyond p = .01) while group EEK-2 did not. Examination of the post-test means for pooled groups K-2 and 3-6 also indicates significant differences, the post mean for the K-2 group being 24.50 and the variance 38.35; and the post-test mean for the 3-6 group being 27.77 and the variances 29.18. The t value for the difference in means is 2.47, significant at p = .05.

In interpreting these data a number of assertions are possible. First there is high probability that arithmetic methods and student teaching in grades 3-6, taken together, contribute significantly to performance. Second, the finding that there is a significant difference between the post-test means of the

3-6 group and the K-2 group suggests that there are differential treatments in student teaching according to grade level, such that experience in teaching in grades 3-6 leads to an increase in performance while experience in teaching in grades K-2 does not. Further evidence on this point with respect to teachers of grades K-3 and grades 4-6 appears on later pages. Third, observation of the fact that the variances of the groups that show significant mean gains decrease, while the variance of the group (EEK-2) that shows no significant mean gain remains stable, strongly suggests that the gains in those groups which show gains are not equally distributed among the members. An increase in mean score and a decrease in variance clearly implies that the initially "weaker" members of the groups gained disproportionately more than the initially "stronger" members. Thus the general effect of arithmetic methods and student teaching is to make the group more homogeneous and at the same time increase the general level of performance--a rather peculiar finding.

Intelligence, Reading Comprehension, and Arithmetic Achievement. Intelligence, reading comprehension, and arithmetic achievement were examined by finding the r of each variable with pre-test scores, gain scores and post-test scores. In order to maintain clarity with respect to correlations between each of the independent variables in question and Form E and Form F of the problems, the undergraduate group was separated on the basis of whether Form E or Form F was taken for the pre-test. Separating the total group on this criterion alone necessarily resulted in pooling subjects who did student teaching in grades K-2 with those who did it in grades 3-6. ACE, Coop C_2,

and STEP Scores were not available for all subjects. Results may be observed in Table 3.

TABLE 3. CORRELATIONS BETWEEN ACE, COOP C_2 AND STEP SCORES AND THE PRE-TEST, POST-TEST, AND GAINS SCORES OF STUDENT TEACHERS.

Group	Score	Form		Test					
		Pre	Post	n	ACE	n	Coop C_2	n	STEP
K-6	Pre	E	E	34	.32*	34	.45***	41	.28*
K-6	Gain	E	E	34	-.22	34	-.21	41	.09
K-6	Post	E	E	34	.29*	34	.36**	41	.41***
K-6	Pre	F	E	28	.40**	28	.45**	34	.35**
K-6	Gain	F	E	28	.00	28	-.32*	34	-.17
K-6	Post	F	E	28	.45**	28	.19	34	.27

* significant at p = .10
** significant at p = .05
*** significant at p = .01

Observations of the r's in Table 3 suggest that none of the independent variables in question can be construed as good predictors of post-test performance, although the fact that the ACE holds small but consistent relationships to performance should not be over looked. The tendency of gains scores to be negatively correlated with test scores is in line with finding that the initially weaker students tend to show the greater gains. It might be noted in this regard that the scattergrams (not shown) between pre-test scores and gain scores also indicated a small negative relationship between these two dependent variables.

MTAI and Study of Values. Product moment correlations were computed between MTAI Scores, each scale of the Study of Values and post-test scores. Since whether the student had previously been tested on either Form F or Form E is not focal in these relationships, all students who took the post-test and returned score sheets for the MTAI and Study of Values were included. This group was divided according to whether student teaching had been done in grades K-2 or grades 3-6. Results may be observed in Table 4.

TABLE 4. CORRELATIONS BETWEEN MTAI AND STUDY OF VALUES SCORES, AND THE POST-TEST SCORES ON STUDENT TEACHERS.

Group	Scale							
	n	MTAI	n	Theo.	Econ.	Aesth.	Soc.	Polit.	Relig.
K-2	30	.03	31	.41*	.06	.06	.44*	-.07	-.18
3-6	29	-.25	32	.10	-.19	.23	.20	-.14	-.41*

* significant at p = .05

It is apparent that there is no relationship between MTAI scores and performance, a finding at variance with those previously reported by Turner (9). However, the inconsistent relationships between Study of Values scales and performance are in agreement with previous findings.

Variables Bearing on Teacher Performance

Grade level. The inclusion of grade 3 teachers in the sample under

consideration made implicit the hypothesis that these teachers do not differ in performance from teachers of grades 4-6. This hypothesis was, of course, verified in a previous study, but the modification of the problems used between the studies necessarily occasioned re-examination of it. In testing the hypothesis, the performance of teachers who were teaching grade 3 was compared by the t test to the performance of teachers of grades 4-6. Largely as a matter of happenstance, none of the 3rd grade teachers in the sample had taught beyond this level, making a clearer delineation between grade levels than was possible in the previous study (9). Results of the comparison may be observed in Table 5.

TABLE 5. PERFORMANCE OF TEACHERS OF GRADE 3 VERSUS PERFORMANCE OF TEACHERS OF GRADES 4-6.

Group	n	s^2	$\bar{X}$	t	p
Teachers Grade 3	32	60.49	24.344	3.332	.01
Teachers Grade 4-6	68	32.22	29.529		

The results shown in Table 5 indicate that teachers of grade 3 are at a distinct disadvantage in solving the problems, and that the problems should not be considered suitable for this group. Because of this finding, grade 3 teachers were deleted in subsequent comparisons.*

*The necessity to eliminate grade 3 teachers was unexpected, and reduced the total N for teachers much below that provided in the original sample. Hence, it is necessary for the reader to note the consistency of the findings reported in later pages, rather than placing great stress on the findings given in individual Tables.

The findings that experienced teachers of grade 3 do not perform as well as experienced teachers of grades 4-6 raises a question concerning whether student teachers with grade 3 experience underperform student teachers of grades 4-6, and should therefore have been grouped with K-2 student teachers rather than those experienced in grades 4-6 in the earlier comparisons. The pre-test and post-test means for student teachers of grade 3 and student teachers of grades 4-6 may be observed in Table 6.

TABLE 6. PRE-TEST AND POST-TEST MEANS OF STUDENT TEACHERS OF GRADE 3 AND STUDENT TEACHERS OF GRADES 4-6.

Group	n	Pre-test $\bar{X}$	Post-test $\bar{X}$
Student teachers Grade 3	9	21.330	28.000
Student teachers Grades 4-6	30	19.830	27.700

On the basis of Table 6, it is apparent that there is no evidence that student teachers of grade 3 underperform student teachers of grades 4-6. This finding suggests the hypothesis that the differences between grade 3 teachers and teachers of grades 4-6 develop as a function of teaching experience occuring subsequent to student teaching.

Teaching situation and preparatory institution. Because not only the separate effects of these variables but also the interaction between them is of interest, both variables were dichotomized for a 2 x 2 analysis of variance. Preparatory institutions were divided by type, i.e. public or private, and

teaching situations according to the size of city in which the teacher taught, with teachers from cities of 24,999 or less (small cities) in one group and teachers from cities of 25,000 or more (large cities) in the other.* It should be noted that the teachers in the small city group taught in non-suburban cities in non-consolidated school systems. The means and variances of each group and the results of the analysis may be observed in Tables 7 and 8 respectively. Three teachers who did not hold college degrees were deleted from the analysis in order to avoid introducing a third variable.

TABLE 7. MEANS AND VARIANCES OF TEACHERS GRADES 4-6 PUBLIC VERSUS PRIVATE COLLEGES, SMALL VERSUS LARGE CITIES.

Group	n	s^2	$\bar{X}$
a. Private College, Small city	11	20.26	26.364
b. Private College, Large city	22	32.37	29.227
c. Public College, Small city	14	33.49	28.429
d. Public College, Large city	18	32.38	33.833

The F value for city is significant beyond p = .01 and for college beyond p = .05. The F for interaction is not significant.

* All teachers who taught in small cities and were prepared in private colleges, were prepared at private colleges of less than 1000 students.

TABLE 8. ANALYSIS OF VARIANCE CITY SIZE X TYPE OF TEACHER PREPARATORY INSTITUTION.

Source	df	unadjusted ss*	adjusted ss	Mean Square
City	1	219.86	261.67	261.67
College	1	165.95	207.76	207.76
Interaction	1	70.34	28.53	28.53
Between	3	456.15		
Within	61	1868.35		30.63
Total	64			

* The adjustment term for disproportionality is 41.81; it is added to the ss for main effects and subtracted from the ss for interaction (11, p.213).

Teaching experience. In order to determine the effects of teaching experience it is necessary to control both source of preparation and type of teaching situation and to begin at zero years of experience. While the teacher sample was initially drawn with a view toward observing the effects of experience from 0-10 years, the unsuitability of the problems for teachers of grade 3 and the necessity for controlling the above variables necessitated a change in the experience groupings. To hold the number of subjects in each group above 5, only four experience categories were used: 0 years (student teachers) 1-2 years, 3-5 years, and 6-10 years.

In the first comparison, teachers were divided according to type of

preparatory institution (private and public) and by experience groups, with type of teaching situation held approximately proportional between experience groups. Since all experienced teachers appearing in these groups taught between grades 4 and 6, the student teacher (0 experience) group was restricted to those students who took the post-test and who either did student teaching in grades 4-6 or else student taught in grades K-3, but were subsequently placed in grades 4-6 for their first full time teaching position. This procedure insured reasonable comparability between groups. Results may be observed in Tables 9 and 10.

TABLE 9. MEANS AND VARIANCES OF TEACHERS WITH 0, 1-2, 3-5, AND 6-10 YEARS OF TEACHING EXPERIENCE PREPARED AT PUBLIC INSTITUTIONS.

Group	n	s^2	$\bar{X}$
0 yrs. experience	35	31.78	27.514
1- 2 yrs. experience	12	41.27	32.000
3- 5 yrs. experience	9	26.75	33.333
6-10 yrs. experience	11	47.25	29.363

With respect to the mean values appearing in Table 9, analysis of variance yielded an F value of 3.18 significant at p = .05, a t test between the 0 years group and the 1-2 years group yielded a t value of 7.23, significant at p = .01, suggesting that the very first years of teaching experience contribute significantly to performance. The drop in performance with 6-10 years experience is attributable to the presence of 3 teachers from small cities who

performed very poorly. Control for city size in a subsequent comparison (p. 33) removed this effect.

TABLE 10. MEANS AND VARIANCES OF TEACHERS WITH 0 YEARS OF EXPERIENCE PREPARED AT A PUBLIC INSTITUTION AND 1-2, 3-5, AND 6-10 YEARS EXPERIENCE PREPARED AT PRIVATE INSTITUTIONS.

Group	n	s^2	$\bar{X}$
0 yrs. experience	35	31.78	27.514
1- 2 yrs. experience	10	25.29	25.800
3- 5 yrs. experience	9	26.20	28.778
6-10 yrs. experience	14	31.76	29.714

Comparison of the means arrayed in Table 10 by analysis of variance yielded an F value of 1.15, which was not significant. However, because the zero experience group was prepared at a public institution, this result hardly demonstrates that teachers from private institutions do not gain with experience, but indicates simply that the performance of teachers prepared at private institutions does not differ significantly from that of undergraduates at a public institution, regardless of years of experience up to 10. This result, it may be noted, is consistent with previous findings (9).

In a second examination of performance as a function of teaching experience, teachers were divided by type of teaching situation and then by experience groups (0, 1-2, 3-5, and 6-10 years of experience). These groups were initially greatly disproportional with respect to the type of institution at

which their members were prepared. In order to achieve proportionality among the groups of experienced teachers, subjects were randomly discarded from each group until exact proportionality was achieved. The student teacher (0 years experience) group could not, of course, be made proportional with respect to college preparation since all graduated from a public institution. However, not all students who did student teaching in grades 4-6 or who were subsequently placed in grades 4-6 took jobs in large cities. Some took jobs in small cities and some were not yet placed at the time the data were run. The latter students were deleted in the comparisons. Results may be observed in Tables 11 and 12.

TABLE 11. MEANS AND VARIANCES OF TEACHERS WITH 0, 1-2, 3-5, AND 6-10 YEARS OF TEACHING EXPERIENCE TEACHING IN CITIES 25,000 AND ABOVE.

Group	n	s^2	$\bar{X}$
0 yrs. experience	25	20.61	26.880
1- 2 yrs. experience	10	54.40	30.800
3- 5 yrs. experience	10	44.62	32.200
6-10 yrs. experience	14	17.76	33.070

Comparison of the means in Table 11 by analysis of variance yielded an F value of 4.87, significant at $p = .05$. A t test between the 0 years group and 1-2 years group was not significant, but differences between the 0 years group and each of the remaining groups was significant beyond $p = .01$.

TABLE 12. MEANS AND VARIANCES OF TEACHERS WITH 0 YEARS OF EXPERIENCE TEACHING IN CITIES 25,000 AND ABOVE AND TEACHERS WITH 1-2, 3-5, AND 6-10 YEARS OF EXPERIENCE TEACHING IN CITIES 24,999 AND BELOW.

Group	n	s^2	$\bar{X}$
0 yrs. experience	25	20.61	26.880
1- 2 yrs. experience	6	17.47	28.333
3- 5 yrs. experience	6	13.37	29.833
6-10 yrs. experience	8	36.78	24.750

Comparison of the means in Table 12 by analysis of variance yielded an F value of 1.49, which was not significant. As can be observed in Table 12, the decline of the means for the 6-10 year group as shown in Table 9 is attributable to teachers prepared at public institutions who have remained in small cities past the fifth year of experience.

A consideration quite relevant to the finding that performance tends to increase as years of experience in grades 4-6 increases, is whether teachers of grades K-3 show corresponding increases with experience. Insofar as the problems in teaching arithmetic would not be expected to be relevant for the latter group, they may be seen as essentially performing the function of controls. Differences between 4-6 teachers and K-3 teachers, all of whom graduated from public colleges, may be observed in Table 13. The groups are approximately proportional with respect to city size.

While the number of observations in each group shown in Table 13 is for the most part small, the consistency of the differences between groups as

experience increases suggests that teaching in grades K-3 is not only much less relevant to performance than teaching in grades 4-6, but that teaching in the primary grades may create negative, rather than positive transfer to solving certain kinds of problems in the intermediate grades.

TABLE 13. TEACHERS OF GRADES 4-6 VERSUS TEACHERS OF GRADES K-3 AT FOUR LEVELS OF TEACHING EXPERIENCE.*

Years of Experience	Grade Level	n	s^2	$\bar{X}$	t	p
0	K-3	56	44.73	26.071	1.06	>.20
	4-6	35	31.78	27.514		
1- 2	K-3	11	80.60	24.818	2.221	.05
	4-6	12	41.27	32.000		
3- 5	K-3	7	39.50	21.714	4.974	.001
	4-6	9	26.75	33.333		
6-10	K-3	7	80.33	22.428	1.856	.10
	4-6	11	47.25	29.363		

* A two-tailed t with pooled variances was used.

Age. To investigate the effects of age on performance, teachers with 1-10 years of experience were divided into three age groups: 20-29, 30-39, and 40-49. To achieve effective control over type of teaching situations, it was necessary to eliminate all teachers from small cities. This procedure

left only teachers from large private and large public preparatory institutions (i.e. institutions with more than 1,000 students) in the 40+ group. To make the other two groups as comparable as possible to the 40+ group, teachers from small institutions were also eliminated from these groups. These procedures left the 20-29 and the 40-49 group proportional with respect to graduates of large public and large private institutions, but created a slight imbalance in the 30-39 group which could not be corrected. In addition, it should be noted that the groups were approximately proportional with respect to the extent that each experience group, 1-2, 3-5, and 6- 10 years, is represented in each age group. Results may be observed in Table 14.

TABLE 14. MEANS AND VARIANCES OF TEACHERS 20-29, 30-39, AND 40+ YEARS OF AGE.

Group	n	s^2	$\bar{X}$
20-29	9	52.25	31.000
30-39	8	26.98	30.125
40+	9	48.61	31.111

Comparison of the means arrayed in Table 14 by analysis of variance yielded an F of .06, indicating that mean performance across age groups, when other variables are reasonably well controlled, is quite homogeneous.

MTAI and Study of Values. The relationship between the MTAI scores, each scale of the Study of Values and performance was investigated by computing r's for the sample of small city teachers and the sample of large city

teachers separately. Results may be observed in Table 15.

TABLE 15. CORRELATIONS BETWEEN MTAI SCORES, STUDY OF VALUES SCALES, AND THE PROBLEM SOLVING PERFORMANCE OF TEACHERS GRADES 4-6.

Group	n	MTAI	Theo.	Econ.	Aesth.	Soc.	Pol.	Rel.
Small city Teachers Grades 4-6	21	.06	.10	.01	-.42	.12	.03	.16
Large city Teachers Grades 4-6	27	.00	-19	.20	.48**	-.21	.44**	.00

** Significant at p = .05

Observation of Table 15 indicates that the MTAI is consistently unrelated to problem solving performance among teachers, a finding consistent with those for student teachers. The relationships between the various scales of the Study of Values and performance among teachers, like those among student teachers, are inconsistent, the larger r's appearing to be random phenomena.

Interproblem Relationships

The interrelationships between problems, and between each problem and the total score, were computed for all teachers of grades 4-6, including the three who did not hold degrees, and for all student teachers grades 4-6 together with those students who student taught in grades K-2 but took teaching positions in grade 6. The intercorrelation matrices for teachers and for student teachers may be observed in Tables 16 and 17, respectively.

TABLE 16. INTERPROBLEM RELATIONSHIPS FOR 69 TEACHERS GRADES 4-6.

Problem	2	3	4	5	Total Score
1	.13	-.04	.15	.00	.38**
2		.07	-.22	.11	.22
3			-.01	-.08	.24*
4				.22	.75**
5					.61**

* significant at p = .05
** significant at p = .01

TABLE 17. INTERPROBLEM RELATIONSHIPS FOR 47 STUDENT TEACHERS GRADES 3-6.

Problem	2	3	4	5	Total Score
1	.18	.09	-.06	-.05	.27
2		.32*	.31*	-.13	.50**
3			.08	-.07	.37**
4				-.13	.75**
5					.35*

* significant at p = .05
** significant at p = .01

Inspection of Tables 16 and 17 indicates that performance on one problem tends to be independent of performance on another, although slight

relationships appear to exist between problem 2 and problems 3 and 4 in the student teacher sample. All r's between problems and total score are significant at p = .10, suggesting that each problem contributes to the total score although there are slight differences in the amount of contribution of each problem between the two samples. These differences did not seem of sufficient moment to warrant further intensive investigation.

DISCUSSION AND CONCLUSIONS

From the pattern of results, it is reasonably apparent that those hypotheses stemming from the assumption that measures of values, attitudes or dominant interests, as exemplified by the MTAI and Study of Values, were non confirmed, while those hypotheses stemming from the assumption that performance varies with opportunity to learn were, for the most part, confirmed. The failure to find the MTAI and Study of Values consistently relevant suggests that subsequent examination of personality variables in relation to problem solving might be more profitably carried on if only variables of demonstrated relevance to teaching, such as those recently reported by Ryans (5), are utilized.

The configuration of findings bearing on the assumption that performance varies with opportunity to learn needs careful examination. The finding that performance increases with exposure to arithmetic methods and student teaching is significant to this assumption for two reasons: first, because it indicates that performance is influenced by specifiable learning treatments, and second because it indicates that the differences in the performance of the graduates of different types of preparatory institutions could quite possibly arise on the basis of differential learning treatments in methods courses and student teaching.

The finding that the performance of third grade and K-3 teachers differs from that of teachers of grades 4-6 and the finding that K-2 student teachers differ in performance from 4-6 student teachers, is also significant to the assumption in question, because they strongly suggest that positive transfer

to problem solving occurs from a comparatively narrow range of antecedent teaching experiences in which problems quite similar to those used in the test trials have been encountered. Thus an "opportunity to learn" means in part an opportunity to encounter problems similar to those used in the criterion measure.

The findings which indicate that performance varies by type of preparatory institution, by type of teaching situation and with experience may be interpreted as generally supporting the assumption that performance varies with opportunity to learn. There are,however, certain discrepancies in these findings. One is the failure to obtain a significant interaction between type of institution at which a teacher prepared and type of teaching situation in which he later teaches. While the failure to find a significant interaction does not negate the assumption that performance varies according to opportunity to learn, it is notable that significant main effects without a significant interaction could arise through selection as well as through differential opportunities to learn. For instance, it is possible that teachers prepared at the private institutions included in the sample are slightly less capable than those prepared at public institutions, and that only the more capable of each of these groups enter teaching in large cities while the less capable teach in small cities. Given the moderate correlation between intelligence and performance, these arrangements would produce exactly the results found. A matter which lends credibility to the latter interpretation is the low mean of small city teachers of 6-10 years of experience, as shown in Table 12. This mean probably indicates that the best teachers in small cities are attracted to better positions, undoubtedly in larger cities.

It is possible, of course, that a complex set of relationships exists between opportunities to learn, ability to learn and teacher selection practices in school systems. While such relationships cannot be extracted from the present data, the possibility, indeed the probability, that they exist should not be overlooked in subsequent studies.

REFERENCES

1. Allport, G. W.; Vernon, P. E.; and Lindzey, Gardner, Study of Values 3rd Ed., Houghton Mifflin Company, Boston, 1960.
2. Cook, W. W.; Leeds, C.H.; and Callis, R., Minnesota Teacher Attitude Inventory, Form A., Psychological Corporation, New York, 1951.
3. Davis, C.C. and Davis, T.G., Cooperative English Test, Test C_2, Reading Comprehension, Higher Level, Form Z, Educational Testing Service, Cooperative Test Division, Princeton, New Jersey.
4. Educational Testing Service, Cooperative Test Division, Sequentional Tests of Educational Progress, Form 3A, Mathematics, Educational Testing Service, Princeton, New Jersey, 1957.
5. Ryans, David G. Characteristics of Teachers, American Council on Education, Washington, D. C., 1960.
6. Thurstone, L.L. and Thurstone, T.G., American Council on Education Psychological Examination for College Freshmen, Educational Testing Service, Cooperative Test Division, Princeton, New Jersey, 1952.
7. Turner, Richard L. and Fattu, Nicholas A., Skill in Teaching, A Reappraisal of the Concepts and Strategies in Teacher Effectiveness Research, Bulletin of the School of Education Indiana University, Bloomington, Indiana, Vol. 36, No. 3, May, 1960.

8. Turner Richard L. and Fattu, Nicholas A., Problem Solving Proficiency Among Elementary School Teachers 1. The Development of Criteria, Monograph of the Institute of Educational Research at Indiana University, Bloomington, Indiana, May, 1960.

9. Turner Richard L., Problem Solving Proficiency Among Elementary School Teachers II. Teachers of Arithmetic, Grades 3-6, Monograph of the Institute of Educational Research at Indiana University, September, 1960.

10. Wade, Eugene W., Problem Solving Proficiency Among Elementary School Teachers III. Teachers of Reading, Grades 2-5, Monograph of the Institute of Educational Research at Indiana University, Bloomington, Indiana, (In Press).

11. Wirt, J.E.; Neidt, C.O.; and Ahman, J.S., Statistical Methods in Educational and Psychological Research, Appleton-Century-Crafts, New York, 1954.